

# Two Approaches to Metaphor Detection

Brian MacWhinney and Davida Fromm

Carnegie Mellon University

Pittsburgh PA, 15213 USA

E-mail: macw@cmu.edu, fromm@andrew.cmu.edu

## Abstract

Methods for automatic detection and interpretation of metaphors have focused on analysis and utilization of the ways in which metaphors violate selectional preferences (Martin, 2006). Detection and interpretation processes that rely on this method can achieve wide coverage and may be able to detect some novel metaphors. However, they are prone to high false alarm rates, often arising from imprecision in parsing and supporting ontological and lexical resources. An alternative approach to metaphor detection emphasizes the fact that many metaphors become conventionalized collocations, while still preserving their active metaphorical status. Given a large enough corpus for a given language, it is possible to use tools like SketchEngine (Kilgariff, Rychly, Smrz, & Tugwell, 2004) to locate these high frequency metaphors for a given target domain. In this paper, we examine the application of these two approaches and discuss their relative strengths and weaknesses for metaphors in the target domain of *economic inequality* in English, Spanish, Farsi, and Russian.

Keywords: metaphor, corpora, grammatical relations, selectional restrictions, collocations

## 1. The Task

In our project, code-named METAL, the task was to automatically detect metaphors in the target domain (Kövecses, 2002) of *economic inequality* – a subject of particular current interest in the wake of the financial collapse of 2008 and growing income disparity within and between countries caused by globalization and automation. Within this target domain, we defined the four subdomains of *poverty*, *wealth*, *taxes*, and *social class*. The task of the system was to take short passages discussing this issue, and to identify within each passage the metaphor(s) dealing with one of the target subdomains of economic inequality, if there was one. Once a sentence with a metaphor was detected, the system had to delineate the specific words involved in the metaphor. Finally, the system had to assign an interpretation to the metaphor by specifying the relevant subsegment of the *economic inequality* target domain (poverty, wealth, taxes, or social class), as well as the source domain. The source domain was characterized in terms of conceptual metaphor theory (Lakoff & Johnson, 1980) involving sources such as PATH, CONTAINER, STRUCTURE, and BODY.

The system targeted detection of metaphors in English, Farsi, Russian, and Spanish. All passages in the system were processed using the TurboParser (Martins, Smith, Xing, Aguiar, & Figueiredo, 2010) which produced dependency relations in CoNNL format (Kübler, McDonald, & Nivre, 2009). From these parses, we extracted pairs of words that represented specific grammatical relations (GRs) or “checkables” that could then be processed for violations of selectional restrictions or preferences (Peters & Wilks, 2010; Wilks, 1978). In many cases the checkable included a word from the source domain linked to a word from the target domain. The possible checkables included:

- Action\_Has\_Subject: verb-subject pairs.  
(poverty vanishes)
- Action\_Has\_Object: verb-object pairs.  
(poverty traps)

- Action\_Has\_Modifier: verb-adverb pairs.  
(trap softly)
- Entity\_Has\_Modifier: noun-adjective pairs.  
(grinding poverty)
- Entity\_Entity: noun-noun pairs.  
(wealth ladder)
- Entity\_Has\_Entity: noun-noun pairs for possession.  
(poverty’s trap, burden of poverty)
- Entity\_Is\_Entity: noun-noun pairs for identity.  
(money is the lifeblood)
- Entity\_Prep\_Entity: noun-(prep)-noun pairs.  
(path\_to wealth)

Once a checkable was detected in the parse tree, it was then sent on to four methods for evaluating selectional restrictions or preferences. As an example, consider the sentence: *There are many hyenas feeding on DBKL’s money*. In the relevant GR, *feeding* is the action and *money* is the object in the Action\_Has\_Object relation. Because one does not literally feed on money, this pair is judged to violate a selectional restriction and hence be a likely candidate for detection as a metaphor.

## 2. Selectional Restriction Methods

The first selectional restriction method used the Common Semantic Features (CSF) method (Tsvetkov, Boytsov, Gershman, Nyberg, & Dyer, 2014). For this system, a classifier for metaphor detection was trained on English samples and then applied to each of the four target languages. The CSF classifier used coarse-grained semantic features derived from WordNet (Fellbaum, 1998), psycholinguistic features for abstractness from the MRC database (Wilson, 1988) and Word Representations in Global Context (Huang, Socher, Manning, & Ng., 2012). Training relied on the TroFi example base (Birke & Sarkar, 2007) as parsed into dependency relations by the TurboParser. Because the WordNet and other materials were based on English, this method treated all four languages as roughly equivalent.

The second selectional restriction method, which we will call the TRIPS method, attempted to extract additional information from WordNet to achieve detection through

preference violation (Wilks, 1978). The ontological classification system of WordNet was further elaborated by ontological relations deriving from the TRIPS system (Allen, Swift, & de Beaumont, 2008) and the method of selecting likely figurative synsets from WordNet (Peters & Wilks, 2010). The TRIPS lexicon provides information on semantic roles and selectional restrictions, and the parser can construct the required semantic structures from WordNet entries. TRIPS has been shown to be successful at parsing WordNet glosses in order to build commonsense knowledge bases (Allen et al., 2011). Consider how the system decides how to handle a word like *cooperate* that is not in the TRIPS hierarchy. Within WordNet, it tries unsuccessfully to process the first synset about *collaborate*, but then moves to the second synset which provides the item *work* and then maps this to the TRIPS ontology to derive the correct selectional restrictions.

The third selectional restriction method relied on the matching of the lexical frames for verbs from VerbNet (Baker, Filmore, & Cronin, 2003) along with features derived from WordNet. This method provides acceptable precision, but its coverage (recall) is limited to metaphors that include verbs included in VerbNet. Moreover, because VerbNet is based on English verbs, the extensions to the other languages are incomplete.

The fourth selectional restriction method processed WordNet features using the Scone ontological database (<http://cs.cmu.edu/~sef/scone>). The Scone ontological hierarchy establishes “split sets” such as animate/inanimate or tangible/intangible. Given a collocation such as *green ideas*, Scone would require that only tangible things can have color and therefore *green ideas* must be metaphorical.

### 3. Problems with Selectional Restrictions

The four selectional restriction methods relied on WordNet ontology and the method of checking for selectional or preference violations across grammatical relations. Thus, reliable detection of metaphors depended on first being able to use NLP tools in each language for accurate identification of each of the relevant “checkable” relations. Whenever there were errors in this parsing process, the resultant grammatical relations or checkable could be either wrong or missing. The second, and even more severe, problem with this method was that the synsets or readings for a given lemma in WordNet mix together both literal and figurative usages. For example, the word *grinding* will include both the literal meaning of physical grinding and the figurative meaning of oppression, as in *grinding poverty*. The method of selectional restriction analysis needs to assume that the expectations of a given synset are violated in a given grammatical relation. But if the synset is already figurative, then there will be no violation. We explored several possible ways around this problem. One was to assume that the first listed synset was the basic or literal reading. Unfortunately, this was not always the case. Moreover, some words have several literal synsets mixed in with figurative ones. Another method, used by the TRIPS system, attempts to process the WordNet entries to extract further information.

We also explored ways of addressing this problem by harvesting complete collections of figurative language from the Oxford English Dictionary (OED) or the McGraw-Hill Dictionary of English Idioms and Phrasal Verbs (Spears, 2009). Our plan was to use these conventionalized forms as a filter on the results from the selectional methods. However, before we completed implementation of this method, we developed a more precise method, as described in the next section.

### 4. Corpus-based Analysis

After working for nearly two years with the selectional methods, we then implemented an alternative approach to metaphor detection that relied on corpus analysis, rather than detection of selectional restrictions. This method used the *TenTen* corpora for English, Russian and Spanish (Jakubicek, Kilgarriff, Kovar, Rychly, & Suchomel, 2013) available from SketchEngine ([sketchengine.co.uk](http://sketchengine.co.uk)). These corpora are called TenTen corpora, because they contain over 10 billion words each. In addition, they have been lemmatized and tagged for part of speech and grammatical relations. Grammatical relations between the lemmas are constructed using a SketchEngine (SkE) shallow parsing grammar based on a set of regular expressions formulated over sequences of part of speech categories. SkE grammars were already available within SkE for Russian, Spanish, and English. These grammars were already built into the corpora, allowing for immediate creation of WordSketches, as described below. Ivanova et al. discuss how a SkE grammar was built for German and Khokhlova and Zakharov (2010) describe the construction of a SkE grammar for Russian. For Farsi, we built a SkE grammar from scratch, using the methods described in these papers and in the SkE documentation.

For Farsi, no TenTen corpus was available, so we needed to build our own SkE corpus. We did this by using passages from web searches supplied by Shlomo Argamon of IIT and pointers to relevant URLs from Ron Ferguson of SAIC. Combining these materials into a single corpus, we then lemmatized and tagged this new corpus, using the methods described by Feely et al. (2014). Using the SkE documentation, we then wrote a SkE grammar to detect the most important grammatical relations in Farsi. Although the resultant Farsi corpus only contained 474,773,547 words, rather than the 10 billion words of the TenTen corpora, it was much larger than any other available Farsi corpus and proved adequate for our purposes. This new corpus, called TalkBank Persian, is now publicly available to users of the SkE system, and it is linked to the SkE grammar, thereby providing the ability to use the system to construct WordSketches in Farsi.

Using these resources, we then constructed WordSketches for all terms in our target domains in all four languages. The method of selecting terms from the target domain for searches through corpora is much like that developed by Mason (2004). However, because we could rely on further facilities from SkE in terms of the TenTen corpora and WordSketches, the process we used was easier to extend and replicate and achieved higher

recall and precision. We used the WordSketches to locate specific instances of relevant metaphors in each corpus. To explain how this is done, we need to describe what a WordSketch looks like. To construct a WordSketch for *poverty*, a user simply enters the search word *poverty* in the SkE interface and then clicks on the “WordSketch” option, as illustrated in Figure 1.

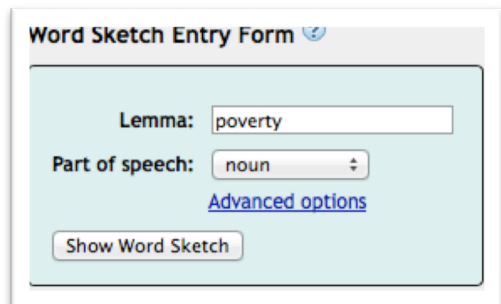


Figure 1: Creating a WordSketch

After clicking on “Show Word Sketch” a large tabular display appears. This table contains columns labeled for each of the grammatical relations (GRs) in which *poverty* appears with matched collocations displayed under each relevant GR label. Figure 2 illustrates a small part of this larger display. For example, it shows *poverty* for the GRs of *object\_of*, *subject\_of*, and *adj\_subject\_of*. Example metaphors include: *object\_of* (*combat poverty*), *subject\_of* (*poverty cripples*), *adj\_subject\_of* (*staggering poverty*), *pp\_obj\_by-i* (*enslaved by poverty*). The full table contains 11 GRs for which there are at least 25 collocations occurring a minimum of 9 times. For the most frequent GRs, setting a threshold of 25 items being displayed means that the lowest frequency items still occur over 100 times. After these top 11 GRs, there are 29 additional GRs with increasingly fewer metaphors. For example, the GR of *part-off-target* only yields 23 cases of the metaphor *shake off poverty* and 9 cases of the non-metaphor *stave off poverty*.

<b>poverty</b> (noun)			enTenTen12 freq = 304221 (23.5 per million) Click on collocates in b		
<b>object_of</b>	<b>59511</b>	<b>0.2</b>	<b>subject_of</b>	<b>29948</b>	<b>0.2</b>
alleviate	3544	9.24	plague	139	4.87
eradicate	2503	9.09	ravage	58	4.82
combat	1272	7.59	beset	40	4.66
halve	330	7.0	blight	32	4.62
grind	861	6.89	afflict	83	4.45
tackle	1527	6.89	persist	95	4.35
perpetuate	311	6.36	exacerbate	62	4.23
escape	1106	6.21	oppress	33	3.81
reduce	7323	6.13	wrack	17	3.71
overcome	1054	6.11	decline	155	3.62
strike	1365	6.01	enslave	23	3.43
			<b>adj_subject_of</b>	<b>3964</b>	<b>0.2</b>
			rampant	119	5.77
			endemic	50	5.44
			rife	55	5.4
			widespread	167	4.27
			pervasive	33	4.27
			prevalent	73	3.67
			first-hand	12	3.13
			staggering	9	2.42
			dire	10	2.12
			inevitable	23	1.92
			evident	38	1.81

Figure 2: First three GRs for POVERTY

For this particular grammatical relation, the English WordSketch returned 304,221 collocations. Each collocation is listed with its frequency of occurrence. For example, *combat poverty* occurs 1,272 times in the corpus, whereas *crush poverty* occurs only 285 times.

Presented with the information constructed automatically by the WordSketch facility, the analyst’s job is to decide which collocations are likely to be metaphorical and which are likely to be literal. This is the one segment of the process that was not automated. However, the process is quite quick, because it is easy to spot in a WordSketch summary which collocations are metaphorical. For example, it is likely that the many cases of *deep poverty* are metaphorical, whereas the instances of *persistent poverty* are not. In some cases, both literal and figurative uses are possible. An example would be *climb the ladder*. Here, we want to distinguish actual physical cases of climbing ladders from metaphorical cases such as *climbing the economic ladder*. To make a closer judgment of the extent to which a given collocation is metaphorical, you can click on the frequency hyperlink next to the collocate, such as *crippling* to access a complete listing of all occurrences of the collocation *crippling poverty* in the TenTen corpus. From the sentence it is possible to click to go further back to the original passage, if necessary. In practice, however, it is easy to glance over the full set of terms in the WordSketch to see which collocations are either transparently metaphorical or transparently literal. This phase of judging metaphoricality only takes perhaps an hour for a given metaphorical target. If the WordSketch facility could trace combinations across additional GRs, these distinctions could be automated. However, given the current technology, such cases simply have to be registered as occasional false alarms.

Using these materials, we then constructed four collections of example metaphors. These include 8,080 sentences for English, 10,263 sentences for Spanish, 6,039 sentences for Russian, and 6,751 sentences for Farsi in the target domain of economic inequality. We could have extracted still more; however, our assigned task was only to locate 6000 sentences for each language. We often had dozens of sentences for a given metaphor. The actual number of distinct conventionalized conceptual metaphors (CCMs) was 963 for English, 866 for Spanish, 428 for Russian, and 911 for Farsi. The four sets of sentences and the four inventories of CCMs have been contributed to Meta-Share. These materials can supplement and extend publicly available materials of the type developed by Shutova and Teufel (2010). Details regarding the construction of these CCM corpora for the four languages and provided in a separate LREC paper (Levin et al., 2014).

## 5. Source Assignment

Part of our task was not just to locate examples of metaphors in the target domain, but also to specify the exact shape of the source domain from which they derive. For example, work with the WordSketch for *poverty* yielded collocations involving words such as *abyss*, *attack*, *combat*, *cure*, *cycle*, *deep*, *defeat*, *disease*, *down*, *edge*, *epidemic*, *escape*, *grind(ing)*, *hell*, *lock*, *ocean*, *pit*, *pressure*, *punish(ing)*, *quagmire*, *sea*, *spread*, *swamp*, *tackle*, *trap*, *verge*, and *wall*. For each of these linguistic metaphors, we needed to assign a conceptual source domain. This work was guided by Zoltán Kövecses, one of the developers of Conceptual Metaphor

Theory (CMT) (Kövecses, 2000, 2002, 2005, 2010; Lakoff & Johnson, 1980). We also relied on the conceptual and grammatical structures provided by the GRs in WordSketches. For example, when deciding how to classify *abyss of poverty*, we needed to consider the source domains of CONFINEMENT and DEPTH. If we invoke CONFINEMENT as the source, we are emphasizing the idea of being trapped in an abyss, much as one is trapped in a prison. On the deepest analytic level, the conceptual metaphor here is PHYSICAL CONFINEMENT IS SOCIAL CONFINEMENT (i.e., the blockage of social movement). If we emphasize DEPTH as source, we are invoking the conceptual metaphor that DOWN IS BAD. In practice, individual linguistic metaphors often mix two, three, or even four conceptual sources. Using the method of spreadsheet-based coding described in the companion LREC paper by Levin et al. (2014), we were then able to code both of these sources as relevant.

## 6. Evaluation

For English, we examined recall and precision for the five methods on a set of 800 utterances that had not been included in the development set. The training set for our system was composed of 400 passages derived from web searches. This test set included 200 passages containing metaphors relevant to economic inequality, 100 with metaphors from other domains, and 100 with no metaphors, as scored by three coders who achieved 94% agreement on coding. The set of 400 training passages was derived from a larger corpus of 2000 passages from which an additional 200 were used for testing.

	Recall	Precision
CSF	.60	.22
TRIPS	.48	.30
VerbNet	.25	.32
Scone	.22	.38
WordSketch	.86	.98

To conduct this analysis, we had to make sure that the metaphors detected by the WordSketch analysis fit into the system of checkables used for the four selectional violation methods. We did this by constructing lexically based regular expression searches for each of the relevant GRs from WordSketch. This additional work was only required to make sure that we could compare across methods. Apart from that, we could have evaluated the results of the WordSketches directly without relying on checkables, because GRs are defined directly in WordSketch grammars in SkE.

The crucial finding of this comparison is that the four methods based on the use of selectional restrictions had high levels of both false negatives and false positives. They frequently confused metonymy with metaphor, identified metaphors outside of the target domain, switched the source and the target, and treated literal terms as metaphorical. These problems could be traced to unclarity in lexical sources, lack of methods for assigning source domain, weakness in the classifiers, errors in parsing, and incomplete coverage through checkables. The results for Farsi, Spanish, and Russian were comparable, although the testing of the selectional

methods for those languages was not completed as a result of our shift at the end of the project to relying on the WordSketch method.

The corpus-based method achieved much better results in terms of both recall and precision. Recall was mostly limited by the fact that not all metaphors in the target domain include target words such as *poverty* or *taxes* clearly linked to the domain. In addition, there are occasionally chains of metaphors that appear partially literal. For example, in *There are many hyenas are feeding on DBKL's money*, there is nothing obviously metaphorical about *hyenas feeding*, until we note that the metaphor of *feeding on money*. The WordSketch method has very high precision. In practice, conventionalized metaphors greatly outnumber their literal alternatives.

The corpus-based method was also clearly superior to the preference violation methods in terms of its ability to identify sources and targets. To do this, we developed methods for extracting metaphors from SketchEngine into GoogleDocs worksheets where we could systematically enter source domain terms for large groups of linguistic metaphors at once. Once this analysis was completed for a given metaphor, it then applied automatically to the hundreds or even thousands of instances of that metaphor in the corpora. In contrast, the preference violation methods had to rely on incomplete and imperfect ontologies to infer source domains from individual sentences, one by one.

These results indicate that, if the goal of the research is to detect metaphors within a well-defined target domain, the corpus-based method is markedly superior to the methods based on selectional restrictions. Application of the corpus-based method requires availability of very large corpora. However, if these are available, the method is quite attractive. Moreover, it is possible to continually refine the method against a constant large corpus. The corpus method also provides a strong foundation for cross-linguistic comparisons. For example, we were able to use this method to track the relative use of metaphors involving POVERTY IS CONFINEMENT across our four languages, showing important differences in frequency of metaphors within this conceptual metaphor across our four languages (see Levin et al, 2014).

However, if the task is to identify metaphors across all domains and false positives are not a concern, then methods based on selectional restrictions may be useful as a supplement to corpus-based methods.

There are non-corpus-based methods that rely on general lexical properties rather than selectional restrictions. It is possible that these methods would perform better than the selectional restriction methods. Turney et al. (2011) describe a method that focuses on the contrast between levels of lexical abstractness in forms such as *shot down the plane* vs. *shot down my argument*. This method can be supplemented by attention to positive valency, as in Turney and Littman (2003). Gandy et al. (2013) describe further extensions of these methods relying on general lexical properties, and Shutova and Korhonen (2012) describe methods based on extensions from basic

metaphor types. However, these methods will all still have to deal with the issue of the mixing of literal and figurative readings within resources such as WordNet.

## 7. Metaphor Distribution

The results of the WordSketch analysis also speak clearly to a second theoretical issue. They show that the vast majority of metaphors available in sources collected from the web are highly conventionalized. For a given domain, such as economic inequality, metaphors related to poverty, education, or taxes use a small number of highly frequent conventionalized patterns or collocations. For example, the 1272 cases of *combat poverty* or the 1106 uses of *escape poverty* show how much we use certain highly conventionalized, but still clearly metaphorical, combinations. In other words, the overwhelming majority of uses of metaphorical language are highly predictable from a simple corpus study.

The fact that nearly all uses of metaphorical collocations are at least partially conventionalized should not obscure the fact that metaphorical language in general can be productive. The detection of individual combinations such as *escape poverty* tells us little about how strings of metaphors can be combined to create larger metaphorical fields in narrative and argumentation (Kövecses, 2010). Moreover, the most creative uses of metaphors are also the rarest. This means that creative metaphors will not be frequent enough to be systematically detected, even in the huge TenTen corpora. None of the methods used so far are able to automatically spot such metaphors. However, it might be possible to use data on metaphorical density to identify particular texts rich in conventionalized metaphors and to then search those carefully by hand to spot truly novel metaphors.

Across the four languages we studied, there are some clear mismatches in the frequency of particular conventionalized metaphors. In previous work, comparisons based on much smaller samples have revealed specific statistical asymmetries. For example, Charteris-Black (2001) compared reporting of economic news in English and Spanish and showed that English uses a greater number of nautical metaphors deriving from its historically greater commitment to sailing. For the domain of economic inequality, some asymmetries of this type arose, such as the greater use of metaphors based on combat in Spanish. However, in the particular domain we studied, the clearest asymmetries involved translational ambiguities. For example, the Spanish word *brecha* can mean both *gap* and *opening* or *path*. The first usage maps to English *abyss*, whereas the second maps more to English *opportunity*. Crosslinguistic comparisons need to be conducted with great care to exclude such simple matters of translational ambiguity.

## 8. Conclusion

Based on the comparison of the numerical results, we can conclude that the WordSketch corpus-based method provides greater recall and precision than the methods based on selectional restriction violations. By comparing comparable corpora across languages, we can understand

both general similarities across cultures and languages and specific divergences, whether they be based on lexical ambiguity, morphological processes, or cultural differences. In addition, the method of analyzing the conceptual bases of these metaphors in spreadsheets with links to the original sources (Levin et al., 2014) produces a rich database for further non-automatic conceptual analysis of the use of metaphors.

On a theoretical level, the most interesting result of this work is that one can construct a nearly complete picture of metaphorical usage for a given target domain by focusing on the detection and analysis of common conventionalized metaphors (CCMs) as revealed through grammatical relations.

## 9. Acknowledgements

This research was supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Defense US Army Research Laboratory contract number W911NF-12-C-0020. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoD/ARL, or the U.S. Government.

## 10. References

- Allen, J., de Beaumont, W., Blaylock, N., Ferguson, G., Orfan, J., & Swift, M. (2011). *Acquiring Commonsense Knowledge for a Cognitive Agent*. Paper presented at the AAAI Fall Symposium Series: Advances in Cognitive Systems (ACS 2011), Arlington, VA.
- Allen, J., Swift, M., & de Beaumont, W. (2008). *Deep Semantic Analysis for Text Processing*. Paper presented at the Symposium on Semantics in Systems for Text Processing (STEP 2008), Venice.
- Baker, C., Filmore, C., & Cronin, B. (2003). The structure of the Framenet database. *International Journal of Lexicography*, 16, 281-296.
- Birke, J., & Sarkar, A. (2007). *Active learning for the identification of nonliteral language*. Paper presented at the Workshop on Computational Approaches to Figurative Language, FigLanguages '07.
- Charteris-Black, J., & Ennis, T. (2001). A comparative study of metaphor in Spanish and English financial reporting. *English for Specific Purposes*, 20, 249-266.
- Feely, W., Manshadi, M., Frederking, R., & Levin, L. (2014). *The CMU METAL Farsi NLP approach*. Paper presented at the LREC 2014, Rejkjavik.
- Fellbaum, C. (Ed.). (1998). *WordNet: An Electronic Lexical Database. Language, Speech and Communication*. Cambridge, MA: MIT Press.

- Gandy, L., Allan, N., Atallah, M., Frieder, O., Howard, N., Kanareykin, S., . . . Argamon, S. (2013). *Automatic identificaition of conceptual metaphors with limited knowledge*. Paper presented at the 27th AAAI conference on Artificial Intelligence.
- Huang, E. H., Socher, R., Manning, C. D., & Ng, A. Y. (2012). *Improving word representations via global context and multiple word prototypes*. Paper presented at the ACL 2012.
- Ivanova, K., Heid, U., Schulte im Walde, S., Kilgariff, A., & Pomikalek, J. (2008). *Evaluating a German Sketch Grammar: a case study on noun phrase case*. Paper presented at the LREC.
- Jakubicek, M., Kilgariff, A., Kovar, V., Rychly, P., & Suchomel, V. (2013). *The TenTen corpus family*. Paper presented at the International Conference on Corpus Linguistics, Lancaster.
- Khokhlova, M., & Zakharov, V. (2010). *Studying Word Sketches for Russian*. Paper presented at the LREC.
- Kilgariff, A., Rychly, P., Smrz, P., & Tugwell, D. (2004). The Sketch Engine *Information Technology Research Institute Reports* (pp. 105-116).
- Kövecses, Z. (2000). *Metaphor and emotion: Language, culture, and body in human feeling*. New York: Cambridge University Press.
- Kövecses, Z. (2002). *Metaphor: a practical introduction*. New York: Oxford University Press.
- Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. New York: Cambridge University Press.
- Kövecses, Z. (2010). A new look at metaphorical creativity in cognitive linguistics. *Cognitive Linguistics, 21*, 655-689.
- Kübler, S., McDonald, R., & Nivre, J. (2009). *Dependency parsing*. San Rafael, CA: Morgan and Claypool.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: Chicago University Press.
- Levin, L., Mitamura, T., MacWhinney, B., Fromm, D., Carbonell, J., Feely, W., . . . Ramirez, C. (2014). *Resources for the detection of conventionalized metaphors in four languages*. Paper presented at the LREC 2014, Rejkjavik.
- Martin, J. H. (2006). A corpus-based analysis of context effects on metaphor comprehension. . In A. Stefanowitsch & S. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 214-236). New York: Mouton de Gruyter.
- Martins, A., Smith, N., Xing, E., Aguiar, P., & Figueiredo, M. (2010). *Turbo parsers: dependency parsing by approximate variational inference*. Paper presented at the EMNLP.
- Mason, Z. (2004). CorMet: A computational, corpus-based conventional metaphor extraction system. *Computational Linguistics, 30*, 23-44.
- Peters, W., & Wilks, Y. (2010). *Data-driven detection of figurative language use in electronic language resources*.
- Shutova, E., & Teufel, S. (2010). *Metaphor Corpus Annotated for Source-Target Domain Mappings*. Paper presented at the LREC.
- Shutova, E., Teufel, S., & Korhonen, A. (2012). Statistical metaphor processing. *Computational Linguistics, 39*, 1-92.
- Spears, R. (2009). *McGraw-Hills dictionary of American idioms and phrasal verbs*. New York: McGraw-Hill.
- Tsvetkov, Y., Boytsov, L., Gershman, A., Nyberg, E., & Dyer, C. (2014). *Metaphor detection with cross-lingual model transfer*. Paper presented at the Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MD.
- Turney, P., & Littman, M. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems, 21*, 315-346.
- Turney, P., Neuman, Y., Assaf, D., & Cohen, Y. (2011). *Literal and metaphorical sense identification through concrete and abstract context*. Paper presented at the Proceedings of the 2011 Conference on the Empirical Methods in Natural Language Processing.
- Wilks, Y. (1978). Making preferences more active. *artificial Intelligence, 11*, 197-223.
- Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, and Computers, 20*, 6-10.